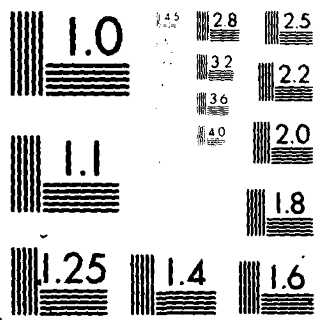
AD-A089 684   BROWN UNIV PROVIDENCE RI LEFSCHETZ CENTER FOR DYNAMI--ETC F/G 12/1
APPROXIMATION SCHEMES FOR NONLINEAR NEUTRAL OPTIMAL CONTROL SYS--ETC(U)
MAY 80 K KUNISCH   DAAG29-79-C-0161

UNCLASSIFIED   AFOSR-TR-80-0925   NL

1 OF 1
AD
A089684

END
DATE
FILMED
10-80
DTIC



MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFOSR-TR-80-0925	2. GOVT ACCESSION NO.: AD-A089 684	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): APPROXIMATION SCHEMES FOR NONLINEAR NEUTRAL OPTIMAL CONTROL SYSTEMS		5. TYPE OF REPORT & PERIOD COVERED: Interim
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): KARL KUNISCH		8. CONTRACT OR GRANT NUMBER(s): DAAG29-79-C-0161, AFOSR 76-3092
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Division of Applied Mathematics, Brown University, Providence, RI 02912		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F 2304/A1
11. CONTROLLING OFFICE NAME AND ADDRESS: Air Force Office of Scientific Research/NM, Bolling AFB, Washington, DC 20332		12. REPORT DATE: May 7, 1980
		13. NUMBER OF PAGES: 54
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): 7 May 80		15. SECURITY CLASS. (of this report): UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

55

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

DTIC ELECTE SEP 29 1980 S D A

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

We discuss methods of approximating stable neutral functional differential equations and associated optimal control problems by sequences of optimal control problems for ordinary differential equations. By introducing a class of "mollified" neutral func. differential equations, convergence of the linear interpolating spline and the averaging approximation scheme is proved. A number of numerical examples is included.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

401 734

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AD A089684

DDC FILE COPY

AFOSR-TR- 80-0925

APPROXIMATION SCHEMES FOR NONLINEAR NEUTRAL OPTIMAL CONTROL SYSTEMS+

by

Karl Kunisch
Institut für Mathematik
Technische Universität Graz
Kopernikusgasse 24
A-8010 Graz, Austria

and

Division of Applied Mathematics
Lefschetz Center for Dynamical Systems
Brown University
Providence, R. I. 02912

May 7, 1980

+This research was supported in part by the Air Force Office of Scientific Research under Contract [redacted]-AFOSR 76-3092[redacted], and in part by the United States Army Research Office under Contract #AROD DAAG29-79-C-0161.

80 9 22 105

APPROXIMATION SCHEMES FOR NONLINEAR NEUTRAL OPTIMAL CONTROL SYSTEMS

by

Karl Kunisch

Abstract:

We discuss methods of approximating stable neutral functional differential equations and associated optimal control problems by sequences of optimal control problems for ordinary differential equations. By introducing a class of "mollified" neutral functional differential equations, convergence of the linear interpolating spline and the averaging approximation scheme is proved. A number of numerical examples is included.

1. Introduction and Notation.

In recent years a significant amount of research was directed towards developing approximation schemes for delay differential equations. Although quite a few attempts have been made before, a paper by Banks and Burns [2] is probably the first one that contains convergence proofs (in a functional analysis setting) for a specific scheme - called averaging approximations - as well as numerical results. Various different schemes, most important spline approximation schemes, have been discussed in the meantime and we refer to [1,2,3,20]. For neutral functional differential equations (NFDE) the author's papers [15] and [7] are probably one of the first contributions to this field. Subsequently in [9] and [11]spline schemes and the averaging approximation scheme have been discussed in different state spaces. For a treatment of a certain class of NFDE's that can be transformed to functional differential equations (FDE) we mention [18].

In this paper we address the question of numerical approximation methods for optimal control problems associated with nonlinear NFDE and we restrict our attention to the averaging or, what is almost equivalent, to interpolating linear spline schemes. Although one might suspect that these schemes converge less quickly than higher order schemes, the extreme simplicity of their algebraic structure and the resulting simplicity in computer programming is a persuasive reason for their investigation. Due to lack of high-order smoothness (averaging approximation subspaces consist, roughly speaking, of L^2 functions and linear interpolating splines are $W^{1,\infty}$ functions), however, it requires much more tedious analysis to verify convergence of these latter schemes than for schemes whose approximating subspaces consist of functions of higher order smoothness [4].

The convergence result that is discussed in this paper generalizes the one in [15] in that it is uniform in the control function that enters into the right hand side of the equation and it extends [9,11] in that these latter papers treat only linear NFDE. The presentation is organized in the following way. Section 2 contains the existence-uniqueness theory that will be relevant for the rest of the paper and introduces the "mollified NFDE". In Section 3 we first discuss an approximation result for nonlinear FDE which generalizes those existing in the literature in that it allows nonlinear point delays and subsequently we apply this result to the mollified NFDE to prove the above-mentioned convergence result for NFDE. Section 4 finally contains a discussion of a class of optimal control problems and examples.

Throughout the paper we shall employ the following notation. The n-dimensional Euclidean space is denoted by $\mathbb{R}^n$ and it is endowed with the Euclidean norm unless specified otherwise. The set of all nonnegative real numbers is $\mathbb{R}^+$. For $-\infty<a<b<\infty$ the space of all continuous functions $x:[a,b] \to \mathbb{R}^n$ is denoted by $C(a,b;\mathbb{R}^n)$ and is endowed with the suprenum norm $\|x\|_{[a,b]}$. We let $L^p(a,b;\mathbb{R}^n)$, $1 < p \leq \infty$, stand for the Banach space of all equivalence classes of functions $x:(a,b) \to \mathbb{R}^n$ such that $|x|^p$ is integrable and denote the usual norm by $|\cdot|_{p,[a,b]}$. There will arise no need to distinguish between representatives and equivalence classes of functions in $L^p(a,b;\mathbb{R}^n)$. The linear space of locally integrable, essentially bounded functions is denoted by $L^{\infty,loc}(a,\infty;\mathbb{R}^n)$ and $W^{1,\infty}(a,b;\mathbb{R}^n)$ stands for the Sobolov space of absolutely continuous functions with derivative in $L^\infty(a,b;\mathbb{R}^n)$. For $C(-r,0;\mathbb{R}^n)$, $r > 0$, we simply write C with norm $|\cdot|$; analogously $L^p = L^p(-r,0;\mathbb{R}^n)$ and $W^{1,\infty} = W^{1,\infty}(-r,0;\mathbb{R}^n)$. We shall also need the space of continuous functions $x:[a,b] \to C$ abbreviated by $C(a,b;C)$ and endowed with suprenum norm $\|\cdot\|_{[a,b]}$. The restriction of a function ϕ to a subset J of its domain is denoted by $\phi|J$. We let $\mathscr{L}_{n,m}$ stand for the vector space of all real $n \times m$ matrices and A^* will be the transpose of $A \in \mathscr{L}_{n,m}$. Finally, as usual in the theory of delay equations, for $r > 0$, $\alpha > 0$ and $x:[-r,\alpha] \to \mathbb{R}^n$ the function x_t is defined by $x_t(s) = x(t+s)$ for $s \in [-r,0]$. Some familiarity with the basic concepts of spline analysis is assumed; as a reference we cite [19]. The reader, who is unfamiliar with delay equations, is referred to [8].

2. Existence Theory and the Mollified Equation.

We consider the nonlinear neutral functional differential equation (NFDE)

$$\left.\begin{aligned} \frac{d}{dt}D(x_t) &= f(t,x_t,u(t)), \quad \text{for} \quad t \geq t_0 \\ x_{t_0} &= \phi, \phi \in C \end{aligned}\right\} \tag{2.1}$$

where

$$D(\psi) = \psi(0) - \sum_{i=1}^{\ell} B_i \psi(-r_i), \quad \text{for} \quad \psi \in C \tag{2.2}$$

with $0 < r_1 \ldots < r_\ell = r$, $B_i \in \mathscr{L}_{n,n}$ and $\mathrm{Dom}(f) = [t_0,\infty) \times C \times \mathbb{R}^m$. A function $x(\cdot;\phi,u)$ or $x(\cdot;\phi)$ or simply x will be called a solution of (2.1) on $[t_0-r,T]$, $T > t_0$, if $x_{t_0} = \phi$ and

$$x(t) = \phi(0) - \sum_{i=1}^{\ell} B_i\phi(-r_i) + \sum_{i=1}^{\ell} B_i x(t-r_i) + \int_{t_0}^{t} f(s,x_s,u(s))ds \tag{2.3}$$

holds for $t \in [t_0,T]$. The following conditions on f will be used:

(H1) For all $\alpha > t_0$, $x \in C(t_0-r,\alpha;\mathbb{R}^n)$ and $u \in L^p(t_0,\alpha;\mathbb{R}^m)$, the mapping $s \to f(s,x_s,u(s))$ from $[t_0,\alpha) \to \mathbb{R}^n$ is integrable;

(H2) for all $\beta > 0$ there exists a nondecreasing function $n_1^\beta \in L^{\infty,loc}(t_0,\infty;\mathbb{R}^+)$ such that

$$|f(t,\phi,u) - f(t,\psi,u)| \leq n_1^\beta(t)(1+|u|^{p/2})|\phi-\psi|$$

for all $t \in [t_0,\infty), \phi$ and $\psi \in C$ with $|\phi| \leq \beta$, $|\psi| \leq \beta$ and $u \in \mathbb{R}^m$;

(H3) there exists a nondecreasing function $n_2 \in L^{\infty,loc}(t_0,\infty;\mathbb{R}^+)$ such that

$$|f(t,\phi,u)| \leq n_2(t)(1+|u|^{p/2})(|\phi|+1),$$

for all $t \in [t_0,\infty)$, $\phi \in C$ and $u \in \mathbb{R}^m$.

Remark 2.1. By a simple calculation one can verify that (H3) will be satisfied, e.g., if (H2) holds and if

(α) there exists a $\tilde{\psi} \in C$ and $n_3 \in L^{\infty,loc}(t_0,\infty;\mathbb{R}^+)$ such that $|f(t,\tilde{\psi},u)| \leq n_3(t)(1+|u|^{p/2})$, and

(β) there exists $\tilde{n}_2 \in L^{\infty,loc}(t_0,\infty;\mathbb{R}^+)$ such that $|f(t,\phi,u)| \leq \tilde{n}_2(t)(1+|u|^{p/2})(|\phi|+1)$ for all $t \in [t_0,\infty)$, $u \in \mathbb{R}^m$ and all $|\phi|$ sufficiently large.

Remark 2.2. Conditions (H1) - (H3) are comparable to conditions (Hi) - (Hiii) in [1]. The conditions in [1] include initial data in $L^2(-r,0;\mathbb{R}^n)$, and linear discrete delays but restrict the class of nonlinear functions f to those arising from distributed delays. The conditions used in this paper include nonlinear discrete delay terms.

Remark 2.3. The most frequently cited class of examples for NFDE is the one arising from certain hyperbolic partial differential equations modelling lossless transmission lines [6,10]. In this case $f = f_1 + f_2$, where f_1 is linear and f_2 is of the form $f_2(\phi) = f_2(\phi(0)+\phi(-r))$ describing the characteristic curve of a diode. So there is a reasonably large class of examples to which our theory applies.

Lemma 2.1. If (H1) - (H3) hold, if $\phi \in C$ and if U is a bounded set in $L^p(t_0,T;\mathbb{R}^m)$ for some $T > t_0$, then

(L1) there exists a unique solution $x(\cdot;\phi,u) \in C(-r+t_0,T;\mathbb{R}^n)$ of (2.1) for each $u \in U$ and

(L2) the family of functions $\{x(\cdot;\phi,u) | u \in U\}$ is a bounded set in $C(t_0-r,T;R^n)$.

We shall not give the proof of this lemma here, since it only involves a standard Piccard-iteration technique. (L2) follows easily from (H3).

We turn to the following "mollified" form of (2.1) and define for $\varepsilon > 0$

$$x(t) = \phi(0) - \sum_{i=1}^{\ell} \frac{1}{\varepsilon} \int_0^{\varepsilon} B_i \phi(s-r_i)ds + \sum_{i=1}^{\ell} \frac{1}{\varepsilon} \int_0^{\varepsilon} B_i x(s-r_i+t)ds + \int_{t_0}^{t} f(s,x_s,u(s))ds. \tag{2.4}$$

Solutions $x^{\varepsilon}(\cdot;\phi,u)$ or $x^{\varepsilon}(\cdot;\phi)$ or simply x^{ε} of (2.4) on $[t_0-r,T)$ are functions satisfying (2.4) with $x^{\varepsilon}_{t_0} = \phi$.

Lemma 2.2. Let $T > t_0, \phi \in C$, let U be a bounded set in $L^p(t_0,T+1;\mathbb{R}^m)$ and assume that (L1), (H1) and (H2) hold. Then

(a) there exists a unique solution $x^{\varepsilon}(\cdot;\phi,u)$ of (2.4) on $[t_0-r,T^{\varepsilon}]$, with $T^{\varepsilon} \geq T$ for all sufficiently small ε and all $u \in U$.

(b) If, moreover, (L2) holds, then

$$\lim_{\varepsilon \downarrow 0} x^{\varepsilon}(t;\phi,u) = x(t;\phi,u)$$

uniformly in $t \in [t_0,T]$ and $u \in U$.

Proof. Parts (a) and (b) will be proved simultaneously. Let $x(\cdot;\phi,u) = x(\cdot)$ be the solution of (2.1) on $[t_0-r,T]$ for some $u \in U$, let $\varepsilon_0 < r_1$ and $\hat{\gamma} = \sup_{u \in U} \|x(\cdot;\phi,u)\|_{[t_0-r,T]}$. For $\alpha \in [t_0,t_0+\min(r_1-\varepsilon_0,T-t_0)]$ we define

$$\mathscr{B}^{\alpha} = \{z \mid z \in C(t_0,\alpha;\mathbb{R}^n),\ \|z\|_{[t_0,\alpha]} \leq \hat{\gamma} + 1\}$$

and for $\varepsilon \in (0,\varepsilon_0)$ and $t \in [t_0,\alpha]$ we let

$$(V^\varepsilon z)(t) = \phi(0) + \frac{1}{\varepsilon}\sum_{i=1}^{\ell}\int_0^\varepsilon B_i(\tilde{\phi}(s+t-t_0-r_i) - \phi(s-r_i))ds + \int_{t_0}^t f(s,\tilde{z}_s,u(s))ds,$$

where $$\tilde{\phi}(s) = \begin{cases} \phi(s) & \text{for } s \in [-r,0] \\ \phi(0) & \text{for } s \in [0,\infty) \end{cases}$$

and $$\tilde{z}(s) = \begin{cases} z(s) & \text{for } s \in [t_0,\alpha) \\ \phi(s-t_0) & \text{for } s \in [t_0-r,t_0] \end{cases}.$$

Obviously $V^\varepsilon(\mathscr{B}^\alpha) \subset C(t_0,\alpha;\mathbb{R}^n)$. We shall verify $V^\varepsilon(\mathscr{B}^\alpha) \subset \mathscr{B}^\alpha$ and that V^ε is a contraction for some $\alpha > t_0$. For $z,w \in \mathscr{B}^\alpha$ and $t \in [t_0,\alpha]$ it follows from (H2) that

$$\begin{aligned}|(V^\varepsilon z)(t) - (V^\varepsilon w)(t)| &\le \int_{t_0}^t |f(s,\tilde{z}_s,u(s)) - f(s,\tilde{w}_s,u(s))|ds \le \\ &\le \|z-w\|_{[t_0,t]}\, n_1^{\hat{\gamma}+1}(T)\int_{t_0}^t (1+|u(s)|^{p/2})ds \le \\ &\le \|z-w\|_{[t_0,t]}\, n_1^{\hat{\gamma}+1}(T)(t-t_0)^{1/2}\Big[(T-t_0)^{1/2} + |u|_{p,[t_0,T]}^{p/2}\Big], \end{aligned} \tag{2.5}$$

where the right hand side of (2.5) is independent of ε, for $0 < \varepsilon < \varepsilon_0$.

We let x be a solution of (2.1) and use (H2) once again to find

$$|(V^{\varepsilon}z)(t)| \leq |x(t)| + |(V^{\varepsilon}z)(t)-x(t)| \leq$$

$$\leq \hat{\gamma} + \underline{|\sum_{i=1}^{\ell} \frac{1}{\varepsilon}\int_0^{\varepsilon} B_i(\tilde{\phi}(s+t-t_0-r_i) - \phi(t-t_0-r_i))ds|} +$$

$$+ |\sum_{i=1}^{\ell} \frac{1}{\varepsilon}\int_0^{\varepsilon} B_i(\phi(s-r_i)-\phi(-r_i))ds| + \int_{t_0}^{t} |f(s,\tilde{z}_s,u(s)) - f(s,x_s,u(s))|ds \leq$$

$$\leq \hat{\gamma} + 2\sum_{i=1}^{\ell} \|B_i\|\, \rho_0(\varepsilon) + 2n_1^{\hat{\gamma}+1}(T)\,(\hat{\gamma}+1)\,(t-t_0)^{1/2}[(T-t_0)^{1/2} + |u|^{p/2}_{p,[t_0,T]}], \tag{2.6}$$

where ρ_0 denotes the modulus of continuity of $x(\cdot)$ on $[-r+t_0,T]$. Estimates (2.5) and (2.6) imply the existence of solutions $x^{\varepsilon}(\cdot;\phi,u)$ on $[t_0-r,\alpha_1]$ for some $\alpha_1 > t_0$ and all $u \in U$. By (H2) and an estimate using the Gronwall lemma, it also follows that

$$\lim_{\varepsilon \downarrow 0} x^{\varepsilon}(t;\phi,u) = x(t;\phi,u) \tag{2.7}$$

uniformly on $[t_0,\alpha_1]$ and uniformly in $u \in U$. Using the uniformities in t and u in inequalities (2.5) and (2.6) we may now proceed stepwise with step-size $\alpha_1 - t_0$ in each step decreasing the range of ε, if necessary, to bound the underlined term in the estimate of $V^{\varepsilon}z$ by an additional use of the triangle inequality and the fact that (2.7) holds on all previous intervals. This concludes the proof. ∎

The last lemma of this section will be of importance for the approximation of optimal control problems associated with (2.1). We shall need an additional hypothesis:

(H4) Conditions (H1)-(H3) hold with $p = 2$ and $f(t,\phi,u) = f_1(t,\phi) + f_2(t,\phi)u$; moreover for all $\alpha > t_0$ and $x \in C(t_0-r,\alpha;\mathbb{R}^n)$, the map $t \to f_2(t,x_t)$ is in $L^2(t_0,\alpha;\mathscr{L}_{n,m})$.

As usual $\rightharpoonup$ will denote weak convergence.

Lemma 2.3. If (H1) - (H4) hold and $u_k \rightharpoonup u$ in $L^2(t_0,T;\mathbb{R}^m)$ for $T > t_0$, then

$$\lim_{k\to 0} x(t;\phi,u^k) = x(t;\phi,u)$$

uniformly in $t \in [t_0,T]$.

Proof. From (H3) it follows that there exists $\hat{\beta} > 0$, such that $|x_t(\cdot;\phi,u)| \le \hat{\beta}$ and $|x_t(\cdot;\phi,u_k)| \le \hat{\beta}$ for all $t \in [t_0,T]$ and $k = 1,2,\ldots$. By (H2) and (H4) we get for $t \in [t_0,\min(T,t_0+r_1)]$

$$|x(t;\phi,u^k) - x(t;\phi,u)| \le \left|\int_{t_0}^{t} (f(s,x_s(\phi,u^k),u^k(s)) - f(s,x_s(\phi,u),u^k(s)))ds\right| +$$

$$+ \left|\int_{t_0}^{t} (f(s,x_s(\phi,u),u^k(s)) - f(s,x_s(\phi,u),u(s)))ds\right| \le$$

$$\leq \int_{t_0}^{t} |f(s,x_s(\phi,u^k),u^k(s)) - f(s,x_s(\phi,u),u^k(s))|ds + |\int_{t_0}^{t} f_2(s,x_s(\phi,u))(u^k(s)-u(s))ds| \leq$$

$$\leq n_1^{\hat{\beta}}(T)[(T-t_0)^{1/2} + (\int_{t_0}^{T} |u^k(s)|^2 ds)^{1/2}](\int_{t_0}^{t} |x_s(\cdot;\phi,u)-x_s(\cdot;\phi,u^k)|^2 ds)^{1/2} + \tilde{\rho}(k),$$

where $\lim_{k\to\infty} \tilde{\rho}(k) = 0$. The last inequality implies

$$|x_t(\cdot;\phi,u^k) - x_t(\cdot;\phi,u)| \leq$$

$$\leq n_1^{\hat{\beta}}(T)[(T-t_0)^{1/2} + |u^k|_{2,[t_0,T]}](\int_{t_0}^{t} |x_s(\cdot;\phi,u)-x_s(\cdot;\phi,u^k)|^2 ds)^{1/2} + \tilde{\rho}(k),$$

so that by an application of a generalized Gronwall lemma and since $\{u^k | k=1,2...\}$ is a bounded subset of $L^2(t_0,T;\mathbb{R}^m)$, the result holds on $[t_0,\min(T,t_0+r_1)]$. Again in a finite number of steps we reach T.

3. An Approximation Result that is uniform in the Control Variable.

In this section we shall prove convergence of an approximation scheme for (2.1) which is uniform in u, as u varies over a bounded set U in $L^p(t_0,T;\mathbb{R}^m)$. This result will then be used in Section 4 to numerically solve optimal control problems associated with (2.1). The idea is to approximate (2.1) by the sequence of mollified equations (2.4), and then to use techniques that have been developed for the approximation of FDE. Of course, passing to the limit, as the FDE converges to the NFDE is the major difficulty that has to be overcome.

We start by considering the FDE

$$\begin{aligned} \dot{\tilde{x}}(t) &= f(t,\tilde{x}_t,u(t)), \quad \text{for} \quad t \geq t_0 \\ \tilde{x}_{t_0} &= \phi, \quad \text{with} \quad \phi \in C. \end{aligned} \tag{3.1}$$

In many instances the reformulation of (3.1) as a Cauchy problem in a function space over the delay interval has proven to be helpful [1,12 et al]. For the space C this has been studied in great detail in [13] and others. It is the variation of constants formula of this abstract Cauchy problem which will be of importance for our purposes. Let $PC = PC(-r,0;\mathbb{R}^n)$ denote the space of piecewise continuous functions $[-r,0] \to \mathbb{R}^n$ endowed with the sup-norm and define $S(t)\colon PC \to PC$ by

$$(S(t)\phi)(s) = \begin{cases} \phi(0) & \text{for} \quad s \geq -t \\ \phi(t+s) & \text{for} \quad -r \leq s \leq -t \end{cases}$$

and $Q_0 : [-r,0] \to PC(-r,0;\mathscr{L}_{n,n})$ by

$$Q_0(s) = \begin{cases} I & \text{for} \quad s = 0 \\ 0 & \text{for} \quad -r \leq s < 0. \end{cases}$$

Consider next the integral equation in C given by

$$\tilde{x}_t = S(t-t_0)\phi + \int_{t_0}^{t} S(t-s)Q_0 f(s,\tilde{x}_s,u(s))ds, \quad \text{for} \quad t \geq t_0 \tag{3.2}$$

where the integral has to be interpreted pointwise as an integral in $\mathbb{R}^n$, i.e.:

$$\left(\int_{t_0}^{t} S(t-s)Q_0 f(s,\tilde{x}_s,u(s))ds\right)(\tau) = \int_{t_0}^{t} \left(S(t-s)Q_0\right)(\tau) f(s,\tilde{x}_s,u(s))ds$$

for $\tau \in [-r,0]$. It is known [12, Proposition 2.1] that for initial data in C and under a condition that is weaker than (H1), (3.1) and (3.2) are equivalent in the sense that $\tilde{x}(t)$ is a solution of (3.1) on $[t_0,T]$ if and only if $\tilde{x}_t$ satisfies (3.2). The integral equation (3.2) and its analogue in the state-space $\mathbb{R}^n \times L^2(-r,0;\mathbb{R}^n)$ have been used to develop various different schemes for FDE before [1,2,4] and Theorem 3.1 below is a generalization of them for a specific class of schemes.

Let $\{t_j^N\}$, $j = 0,\ldots,N$ be a partition of $[-r,0]$ given by $t_j^N = -\frac{r}{N}j$ and define the corresponding sequence of linear finite dimensional subspaces Z_1^N of C by $Z_1^N = \{\phi \in C | \phi$ is a linear spline with knots at $t_j^N\}$.

A basis for Z_1^N is given by the columns of $\beta^N = (\beta_0^N,\ldots,\beta_N^N)$, where β_j^N is a matrix whose columns are in Z_1^N and for which

$$\beta_i^N(t_j^N) = \delta_{ij} I \tag{3.3}$$

holds; here δ_{ij} is the Kronecker symbol and I is the identity matrix. Of course, $\dim(Z_1^N) = N+1$. Next we introduce some additional notation. The families of operators $\{P_1^N\}$ and $\{A_1^N\}$ from C into Z_1^N and $\{Q_1^N\}$ from $[-r,0]$ into $\mathscr{L}_{n,n}$ are defined by

$$(P_1^N\phi)(t_j^N) = \phi(t_j^N), \quad \text{for } j = 0,\ldots,N$$

$$Q_1^N(s) = \begin{cases} 0 & \text{for } s \in [-r,-t_1^N] \\ (1+\frac{N}{r}s)I & \text{for } s \in (-t_1^N,0] \end{cases},$$

and

$$A_1^N\phi = \psi^N ,$$

with $\psi^N \in Z_1^N$ given by

$$\begin{cases} \psi^N(t_j^N) = \frac{N}{r}(\phi(t_{j-1}^N) - \phi(t_j^N)) & \text{for } j = 1,\ldots,N \\ \psi^N(0) = 0. \end{cases}$$

P_1^N and Q_1^N act as interpolation operators onto, respectively, into Z_1^N. We call $\{Z_1^N, P_1^N, A_1^N\}$ the linear interpolating spline scheme [4]. To motivate the definition of A_1^N we recall that $S(t): C \to C$ is a linear C_o-semigroup, whose infinitesimal generator, $\tilde{A}$, is given by $\mathrm{Dom}(\tilde{A}) = \{\phi \mid \phi \in C^1(-r,0;\mathbb{R}^n), \dot{\phi}(0) = 0\}$ and $\tilde{A}\phi = \dot{\phi}$. Since A_1^N, $N = 1,2,\ldots,$ are bounded linear operators, they also generate linear semigroups $e^{A_1^N t}$ that we denote by $S_1^N(t)$. The matrix representation $[A_1^N]$ of A_1^N restricted to Z_1^N is given by

$$[A_1^N] = \begin{pmatrix} \begin{matrix} 0 \\ I \end{matrix} & \begin{matrix} 0 \\ -I \end{matrix} & & 0 \\ & \ddots & \ddots & \\ 0 & & I & -I \end{pmatrix} \otimes \frac{N}{r}, \tag{3.4}$$

where $\otimes$ denotes the Kronecker product.

<u>Remark 3.1.</u> The proofs to the theorems in this section rely quite heavily on the simple structure of $[A_1^N]$ and $e^{[A_1^N]t}$. For higher order splines, for example, the matrices analogous to A_1^N become wider and wider band matrices, whose matrix-exponentials seem quite formidable.

In the following lemma we state some of the properties of A_1^N, P_1^N and S_1^N.

Lemma 3.1.

(a) $\lim_N P_1^N \phi = \phi$ for all $\phi \in C$,

(b) $\| P_1^N \| \leq 1$ for all N,

(c) $\lim_N S_1^N(t)\phi = S(t)\phi$, uniformly in t on compact subsets of $[0,\infty)$,

(d) $\| S_1^N(t) \| \leq 1$ for all N,

(e) $\lim_N (S_1^N(t)Q_1^N)(s) = (S(t)Q_0)(s)$, uniformly on compact subsets of $[0,a] \times [-r,0] \setminus \{(t,s) \mid t = -s\}$, for any $a > 0$,

(f) $\| Q_1^N(s) \| \leq 1$ for all $s \in [-r,0]$ and all N.

The proof of Lemma 3.1. is contained in [13, pp.81]. (a),(b) and (f) are straightforward, (c) follows from the Trotter-Kato theorem, (d) holds since the logarithmic norm of $[A_1^N]$ is zero and (e) is proved by using the special structure of $e^{[A_1^N]t}$.

Theorem 3.1. Let (H1) - (H3) hold and assume that U is a bounded subset of $L^p(t_0,T;\mathbb{R}^m)$ and $T > t_0$. Then for all N sufficiently large

$$z(t) = S_1^N(t-t_0)P_1^N\phi + \int_{t_0}^{t} S_1^N(t-s)Q_1^N f(s,z(s),u(s))ds \tag{3.5}$$

has a unique solution $\tilde{z}^N = \tilde{z}^N(\cdot;\phi,u) \in Z_1^N$ on $[t_0,T]$ and

$$\lim_N \tilde{z}^N(t;\phi,u) = \tilde{x}_t(\cdot;\phi,u) \quad \text{in} \quad C \tag{3.6}$$

uniformly in $u \in U$ and $t \in [t_0,T]$.

Proof. The proof of this theorem is contained in [16] and will only be outlined here. First, it was verified in [13, Lemma B.2] that the integral in (3.5) exists as a Bochner integral and that

$$\left[\int_{t_0}^{t} S_1^N(t-s)Q_1^N f(s,y(s),u(s))ds\right](\tau) = \int_{t_0}^{t} (S_1^N(t-s)Q_1^N)(\tau)f(s,y(s),u(s))ds.$$

Since $S_1^N Z_1^N \subset Z_1^N$ and $Q_1^N Z_1^N \subset Z_1^N$ it follows immediately that the trajectory of any solution of (3.5) must lie in Z_1^N.

By Lemma 2.1, (L2), the family $\{\tilde{z}(t,u) \mid t \in [t_0,T] \text{ and } u \in U\}$ is contained in a ball of radius γ, $0 < \gamma < \infty$, in C. For any α, $t_0 < \alpha < T$ we define the set

$$\mathscr{B}_\alpha = \{\psi \in C(t_0,\alpha;C) \mid \sup_{t \in [t_0,\alpha]} |\psi(t)| \leq \gamma + 1\},$$

and a family of operators $V^N : C(t_0,\alpha;C) \to C(t_0,\alpha;C)$ by

$$(V^N y)(t) = S_1^N(t-t_0)P_1^N\phi + \int_{t_0}^{t} S_1^N(t-s)Q_1^N f(s,y(s),u(s))ds. \tag{3.7}$$

By (H2) and Lemma (3.1) (d) one can show that there exists a constant C_1, independent of N and $u \in U$, such that for all $t \in [t_0,\alpha]$ and $y,w \in \mathcal{B}_\alpha$

$$|V^N y(t) - V^N w(t)| \le C_1 \sup_{t \in [t_0,\alpha]} |y(t) - w(t)|(t-t_0)^{1/2}. \quad (3.8)$$

Moreover, there exists a function ρ with $\lim_N \rho(N) = 0$ and a constant C_2, both independent of $u \in U$ and C_2 not depending on N, such that for all $y \in \mathcal{B}_\alpha$ and $t \in [t_0,\alpha]$

$$|V^N y(t)-z(t)| \le \rho(N)+C_2\left(\tau+\lambda^{1/2}+(t-t_0)^{1/2}+\left(\int_{t_0}^{t} |y(s)-z(s)|^2 ds\right)^{1/2}\right) \le$$

$$\le \rho(N) + C_2(\tau+\lambda^{1/2}+2(t-t_0)^{1/2}(1+\gamma)). \quad (3.9)$$

To verify (3.9) one uses (3.2), (3.7), (H2), (H3) and Lemma 3.1.; Lemma 3.1. (e) specifically implies that for each $\tau > 0$ and $\lambda > 0$ there exists an $N_0 = N_0(\tau,\lambda)$ such that $|(S_1^N(t-s)Q_1^N)(\theta) - (S(t-s)Q_0)(\theta)| \le \tau$ for all $N \ge N_0$, $t \in [t_0,T]$, $\theta \in [-r,0]$ and $s \in [t_0,t] - (t-\lambda+\theta,t+\lambda+\theta)$.

Estimates (3.8) and (3.9) imply that there exist $\tau_0 > 0$, $\lambda_0 > 0$ and $\alpha_1, t_0 < \alpha_1 < T$, such that for all N sufficiently large $V^N \mathcal{B}_{\alpha_1} \subset \mathcal{B}_{\alpha_1}$ and V^N is a contraction on $\mathcal{B}_{\alpha_1}$, so that there exist unique solutions $\tilde{z}^N$ of (3.5) on $[t_0,\alpha_1]$. For $t \in [t_0,\alpha_1]$ equality (3.6) follows from the first inequality of (3.9) with $V^N y(t) = y(t) = \tilde{z}^N(t)$.

We now notice that the constants in the estimates (3.8) and (3.9) are actually uniform in $t \in [t_0, T]$. Therefore, one can proceed stepwise with constant stepsize $\alpha_1 - t_0$, each time repeating the above argument until T is reached. This concludes the proof. ▦

Remark 3.2. Theorem 3.1. remains true, if (H3) is replaced by assuming Lemma 2.1 (L2) and

$$|f(t,0,u)| \leq n_2(t)(1+|u|^{p/2}), \quad \text{for all} \quad t \in [t_0,\infty) \quad \text{and } u \in \mathbb{R}^m.$$

For the proof of Theorem 3.1 we only used properties of the $\{Z_1^N, P_1^N, A_1^N\}$-scheme as given in Lemma 3.1., so that this theorem would remain valid for any scheme that satisfies Lemma 3.1.

We now recall the NFDE (2.1)

$$\frac{d}{dt} Dx_t = f(t,x_t,u(t)) \quad \text{for} \quad t \geq t_0$$

$$x_{t_o} = \phi$$

and assume

(H5) (i) $D\phi = \phi(0) - B\phi(-r_1)$ with $0 < r_1 \leq r$

(ii) $\rho(B) < 1$;

here ρ denotes the spectral radius of the matrix $B \in \mathscr{L}_{n,n}$.

Remark 3.3. Condition (H5) is not as strong as it might appear at first, and is directly connected with stable difference operators as discussed in [8, chapter 12.5]. If D is of the specific form chosen in (H5) (i), then it is stable (stable in the delay r_1) if and only if $\rho(B) < 1$. Note, that if $\rho(B) < 1$, then there always exists a norm $|\cdot|$ on $\mathbb{R}^n$ and a subordinate matrix-norm $\|\cdot\|$, such that $\|B\| < 1$. Throughout the rest of this chapter it is assumed that $\mathbb{R}^n$ is endowed with this norm. It should also be mentioned that those NFDE which arise when transforming certain hyperbolic partial differential equations generally satisfy $\rho(B) < 1$, [6].

Remark 3.4. Although the results are stated for the case when D contains only one discrete delay, they are easily generalized to several delays as in (2.2), as long as there exists some norm on $\mathbb{R}^n$ such that $\sum_{i=1}^{\ell} \|B_i\| < 1$ holds for the subordinate matrix norms. We recall that in the scalar case $\sum_{i=1}^{\ell} \|B_i\| < 1$ is necessary and sufficient for D to be stable, [8, pp.291].

For D as in (H5) the mollified equation (2.4) is just the variation of constants formula of

$$\dot{x}(t) = \frac{1}{\varepsilon} B[x(t+\varepsilon-r_1) - x(t-r_1)] + f(t,x_t,u(t))$$

$$x_{t_0} = \phi, \phi \in C. \tag{3.10}$$

By Lemma 2.2 (a) there exists an ε_1 such that (3.10) has a unique solution $x_t^\varepsilon = x^\varepsilon(\cdot;\phi,u)$ for all ε with $0 < \varepsilon < \varepsilon_1 < r_1$. Next for $\varepsilon \in (0,\varepsilon_1]$ consider the sequence of equations

$$z^{N,\varepsilon}(t) = S_1^N(t-t_0)P_1^N\phi + \int_{t_0}^t S_1^N(t-s)Q_1^N f(s,z^{N,\varepsilon}(s),u(s))ds + \\ + \int_{t_0}^t S_1^N(t-s)Q_1^N \frac{1}{\varepsilon} g(z^{N,\varepsilon}(s)(\cdot+\varepsilon)-z^{N,\varepsilon}(s)(\cdot))ds \tag{3.11}$$

where $g(\phi) = B\phi(-r_1)$, (or $g(\phi) = \sum_{i=1}^{\ell} B_i\phi(-r_i)$ in case of multiple discrete delays in D). By Theorem 3.1 there exist solutions $z^{N,\varepsilon}(\cdot;\phi,u)$ of (3.11) for sufficiently large N and

$$\lim_N z^{N,\varepsilon}(t;\phi,u) = x_t^\varepsilon(\cdot;\phi,u). \tag{3.12}$$

The above limit holds uniformly in $t \in [t_0,T]$ and $u \in U$, for each fixed $\varepsilon \in (0,\varepsilon_1]$.

The sequence of approximating problems whose solutions (hopefully) approximate the solution of (2.1) finally is given by

$$z^N(t) = S_1^N(t-t_0)P_1^N\phi + \int_{t_0}^t S_1^N(t-s)Q_1^N f(s,z^N(s),u(s))ds + \\ + \int_{t_0}^t S_1^N(t-s)Q_1^N \frac{N}{r} B(z^N(s)(t^N_{j^N-1}) - z^N(s)(t^N_{j^N}))ds \tag{3.13}$$

This is the appropriate form of using the $\{Z_1^N, P_1^N, A_1^N\}$ scheme for approximating (2.1).

Theorem 3.2. Let $T > t_0$ and let U be a bounded set in $L^p(t_0,T;\mathbb{R}^m)$. Moreover, assume that (H1) - (H3) and (H5) hold and that $\|B\| < 1$ for an appropriate norm on $\mathbb{R}^n$. Then for $\phi \in W^{1,\infty}(-r,0;\mathbb{R}^n)$ solutions $z^N(t;\phi,u)$ of (3.13) exist and

$$\lim_N z^N(t;\phi,u) = x_t(\cdot;\phi,u) \tag{3.15}$$

uniformly in $u \in U$ and $t \in [t_0,T]$.

Proof. First we need some additional notation. For $x \in \mathbb{R}^{n(N+1)}$, with $x = \mathrm{col}(x_0,\ldots,x_N)$ let $\|x\|^{\infty} = \sup_{i=0,\ldots N} |x_i|$ and $\|x\|^{\infty'} = \sup_{i=1,\ldots,N} |x_i|$. Since the trajectories of (3.11) lie entirely in Z_1^N it follows that $z^{N,\varepsilon}(t;\phi,u) = \beta^N v^{N,\varepsilon}(t)$ uniquely defines $v^{N,\varepsilon}(t;\phi,u) \in \mathbb{R}^{n(N+1)}$, for $t \in [t_0,T]$.

The existence of solutions $z^N(t;\phi,u)$ of (3.14) is quite simple to verify and we immediately turn to (3.15) and choose $\eta > 0$ arbitrarily. By Lemma 2.2(b) it follows that there exists an $\varepsilon_2 < \varepsilon_1$ such that

$$|x_t^{\varepsilon}(\cdot;\phi,u) - x_t(\cdot;\phi,u)| \le \eta \tag{3.16}$$

This is the appropriate form of using the $\{Z_1^N, P_1^N, A_1^N\}$ scheme for approximating (2.1).

Theorem 3.2. Let $T > t_0$ and let U be a bounded set in $L^p(t_0,T;\mathbb{R}^m)$. Moreover, assume that (H1) - (H3) and (H5) hold and that $\|B\| < 1$ for an appropriate norm on $\mathbb{R}^n$. Then for $\phi \in W^{1,\infty}(-r,0;\mathbb{R}^n)$ solutions $z^N(t;\phi,u)$ of (3.13) exist and

$$\lim_N z^N(t;\phi,u) = x_t(\cdot;\phi,u) \tag{3.15}$$

uniformly in $u \in U$ and $t \in [t_0,T]$.

Proof. First we need some additional notation. For $x \in \mathbb{R}^{n(N+1)}$, with $x = \mathrm{col}(x_0,\ldots,x_N)$ let $\|x\|^{\infty} = \sup_{i=0,\ldots N} |x_i|$ and $\|x\|^{\infty'} = \sup_{i=1,\ldots,N} |x_i|$. Since the trajectories of (3.11) lie entirely in Z_1^N it follows that $z^{N,\varepsilon}(t;\phi,u) = \beta^N v^{N,\varepsilon}(t)$ uniquely defines $v^{N,\varepsilon}(t;\phi,u) \in \mathbb{R}^{n(N+1)}$, for $t \in [t_0,T]$.

The existence of solutions $z^N(t;\phi,u)$ of (3.14) is quite simple to verify and we immediately turn to (3.15) and choose $\eta > 0$ arbitrarily. By Lemma 2.2(b) it follows that there exists an $\varepsilon_2 < \varepsilon_1$ such that

$$|x_t^{\varepsilon}(\cdot;\phi,u) - x_t(\cdot;\phi,u)| \le \eta \tag{3.16}$$

for all $\varepsilon < \varepsilon_2$, $u \in U$ and $t \in [t_0,T]$. We shall use the simple estimate

$$\|w^N(t)-v^{N,\varepsilon}(t)\|_\infty \leq \sup(\|w^N(t)-v^{N,\varepsilon}(t)\|_{\infty'}, |w_0^N(t)-w_1^N(t)|+\|w^N(t)-v^{N,\varepsilon}(t)\|_{\infty'} +$$

$$+ |v_1^{N,\varepsilon}(t)-v_0^{N,\varepsilon}(t)|).$$

From the two technical Lemmas 3.2 and 3.3 at the end of this section and the above estimate, it follows that there exist constants N_0 and ε_3, $0 < \varepsilon_3 < \varepsilon_2$, such that

$$|z^{N,\varepsilon}(t;\phi,u) - z^N(t;\phi,u)| \leq \eta \tag{3.17}$$

for all $\varepsilon < \varepsilon_3$, $N \geq N_0$, $u \in U$ and $t \in [t_0,T]$. Finally, (3.12), (3.16) and (3.17) imply that for some $N_1 \geq N_0$ $|z^N(t;\phi,u)-x_t(\cdot;\phi,u)| \leq 3\eta$ for all $N \geq N_1$, uniformly in t and u. This ends the proof. ∎

<u>Remark 3.5.</u> The special form of $[A_1^N]$ was used in the proof of Lemma 3.1 (e) and will be used even more stringently in the proofs of Lemmas 3.2 and 3.3 below. For subspaces Z_{av}^N arising from averaging approximations [2] the operators approximating $\tilde{A}$ turn out to have the same matrix representation as A_1^N, when restricted to the finite dimensional subspaces. We briefly discuss

the averaging approximation scheme. The state space is chosen to be $\mathbb{R}^n \times L^2(-r,0;\mathbb{R}^n)$, and will be abbreviated by Z when endowed with its natural inner product and resulting norm. Z^N_{av} are linear subspaces of Z defined by $Z^N_{av} = \{ (\eta,\phi) \mid \eta \in \mathbb{R}^n, \phi = \sum_{j=1}^{N} a_j\chi_j, \ a_j \in \mathbb{R}^n\}$, where χ^N_j is the characteristic function of the interval $[t^N_j, t^N_{j-1})$. The orthogonal projections $P^N_{av}: Z \to Z^N_{av}$ are given by $P^N_{av}(\eta,\phi) = (\eta, \sum_{j=1}^{N} \psi^N_j\chi^N_j)$ with $\psi^N_j = \frac{N}{r}\int_{t^N_j}^{t^N_{j-1}} \phi(s)ds$ for $j = 1,\ldots,N$. The operators corresponding to Q^N_1 are defined by $Q^o_{av}(\eta,\phi) = (\eta,0)$, independently of N and $A^N_{av}: Z \to Z$ are given by $A^N_{av}(\eta, \sum_{j=1}^{N} a_j\chi^N_j) = (0, \frac{N}{r}\sum_{j=1}^{N}(a^N_{j-1}-a^N_j)\chi^N_j)$ with $a_o \overset{def}{=} 0$, and $A^N_{av}z = A^N_{av}P^N_{av}z$ for $z \in Z$. $\{ Z^N_{av}, P^N_{av}, A^N_{av}\}$ is called the averaging approximation scheme.

For linear equations with f of the form

$$\tilde{f}(\eta,\phi,u(t)) = A_0\eta + \sum_{i=1}^{\nu} A_i\phi(-h_i) + \int_{-r}^{0} A(s)\phi(s)ds + Eu(t), \quad (3.18)$$

where $0 = h_0 < h_1 < \ldots < h_\nu = r$, $A_i \in \mathscr{L}_{n,n}$, $E \in \mathscr{L}_{n,m}$, $A(\cdot)$ an L^2 - $n \times n$- matrix-valued function and $u \in L^p(t_0,T;\mathbb{R}^m)$, it was proved in [2] that

$$\lim_N \tilde{z}^N_{av}(t;\eta,\phi,u) = \tilde{z}(t;\eta,\phi,u) = (\tilde{x}(t),\tilde{x}_t) \text{ in } Z \quad (3.19)$$

uniformly in $t \in [t_0,T]$ and in u, as u varies over a bounded subset of $L^p(t_0,T;\mathbb{R}^n)$. Here $\tilde{x}$ satisfies $\frac{d\tilde{x}(t)}{dt} = f(\tilde{x}(t),\tilde{x}_t,u(t))$, $(\tilde{x}(t_0),\tilde{x}_{t_o}) = (\eta,\phi)$, with $\tilde{f}$ as in (3.18), and $\tilde{z}^N_{av}(t) =$ $= (\tilde{w}^N_0(t), \sum_{j=1}^N \tilde{w}^N_j(t)x^N_j)$ where $\tilde{w}^N(t) \in \mathbb{R}^{n(N+1)}$ are the solutions of the ordinary differential equation

$$\begin{cases} \dot{\tilde{w}}^N(t) = [A^N_1]\tilde{w}^N(t) + \mathrm{col}(\tilde{f}(\tilde{z}^N_{av}(t),u(t)),0,\ldots,0) \\ \tilde{w}^N(t_0) = \mathrm{col}(\eta,\phi^N_1,\ldots,\phi^N_N) \quad \text{where} \quad P^N_{av}(\eta,\phi) = (\eta, \sum_{j=1}^N \phi^N_j x^N_j). \end{cases}$$

Once (3.19) is established one can turn to

$$\begin{cases} \dot{x}(t) = B\dot{x}(t-r_1) + \tilde{f}(x(t),x_t,u(t)) \\ (x(t_0),x_{t_o}) = (\phi(0),\phi), \phi \in C. \end{cases} \tag{3.20}$$

Only a few changes have to be made in the proof of Theorem 3.2. (we notice that the technical Lemmas 3.1 and 3.2 remain true, since $|(P^N_{av}(\phi(0),\phi))_i - (P^N_{av}(\phi(0),\phi))_{i+1}| = 0(\frac{1}{N})$ for $i = 0,\ldots,N-1$, and $\phi \in W^{1,\infty}$) before we arrive at:

<u>Proposition 3.1.</u> If $\tilde{f}$ is as in (3.18), if $T > t_0$, $\phi \in W^{1,\infty}$ and if U is a bounded subset of $L^p(t_0,T;\mathbb{R}^m)$, then

$$\lim_N z^N_{av}(t) = (x(t),x_t) \quad \text{in} \quad Z \tag{3.21}$$

uniformly in $u \in U$ and $t \in [t_0,T]$. Here x is a solution of (3.20) and $z^N_{av}(t) = (w^N_{av,0}(t), \sum_{i=1}^N w^N_{av,i}(t)\chi^u_i)$, where w^N_{av} is the solution of

$$\dot{w}^N_{av}(t) = [A^N_1]w^N_{av}(t) + \mathrm{col}(\tilde{f}(z^N_{av}(t),u(t)) + \tfrac{N}{r}B(w^N_{av,j^N-1}(t)-w^N_{av,j^N}(t)),0,\ldots 0)$$

$$w^N(t_0) = \mathrm{col}(\phi(0),\phi^N_1,\ldots,\phi^N_N) \quad \text{where} \quad P^N_{av}(\phi(0),\phi) = (\phi(0),\sum_{i=1}^N \phi^N_i\chi^N_i). \tag{3.22}$$

Comparing (3.14) with (3.22) we find that for many specific NFDE the approximating ordinary differential equations arising from the linear interpolating spline scheme and the averaging approximation scheme differ only with respect to the initial value. The initial values will be equal if ϕ is constant, for example. Proposition 3.1. can be extended along the lines of Theorems 3.1 and 3.2 to include nonlinear equations, if conditions analagous to (H1)-(H3) hold. Since the techniques are rather obvious but tedious, we shall not include the details. This ends Remark 3.5.

We conclude this section with the two lemmas which are necessary for the proof of Theorem 3.2 and recall that $v^{N,\varepsilon}(t)$ and $w^N(t)$ are the coordinate vectors of $z^{N,\varepsilon}(t;\phi,u)$ and $z^N(t;\phi,u)$ with respect to the basis β^N respectively.

Lemma 3.2. Under the hypotheses of Theorem 3.1 there exist constants K_i, $i = 1,\ldots,8$ depending on ϕ, f, B, T and U, but not depending on N, ε, t and $\dot{\phi}$ such that

$$\|w^N(t)\|^\infty \leq K_1 \quad \text{and} \quad \|\dot{w}^N(t)\|^\infty \leq K_2|\dot{\phi}|_\infty + K_3 + K_4|u(t)|^{p/2}. \tag{3.23}$$

If, in addition, $\varepsilon > 0$ and N satisfy $\frac{r}{\varepsilon N} < \frac{1-\|B\|}{2}$ then

$$\|v^{N,\varepsilon}(t)\|^\infty \leq K_5 \quad \text{and} \quad \|\dot{v}^{N,\varepsilon}\|^\infty \leq K_6|\dot{\phi}|_\infty + K_7 + K_8|u(t)|^{p/2} \tag{3.24}$$

with $|\dot{\phi}|_\infty = \operatorname*{ess\,sup}_{t\in[-r,0]} |\dot{\phi}(t)|$.

Proof. We start with a few useful technicalities and define

$$Z(s,i) = \frac{N}{i!r}\left(\frac{N}{r}(t-s)\right)^i \exp\left(-\frac{N}{r}(t-s)\right) \quad \text{for } i = 0,\ldots,N-1 \text{ and } t \geq s.$$

The following estimates can be found:

$$\int_{t_0}^{t} Z(s,i)ds \leq 1 \quad \text{for} \quad i = 0,\ldots,N-1 \tag{3.25}$$

and

$$\int_{t_0}^{t} Z^2(s,i)ds \le \frac{N\, e^{\frac{1}{24i}}}{2r\sqrt{\pi i}}\ , \quad \text{for} \quad i=0,\ldots,N-1. \tag{3.26}$$

(3.25) follows by a simple induction argument showing $\int_{t_0}^{t} Z(s,i)ds \le 1 - \sum_{k=0}^{i} (\frac{1}{i!}) \frac{N(t-t_0)^k}{r} \exp\,(\frac{-N}{r}(t-t_0))$. Another induction argument implies that $\int_{t_0}^{t} Z^2(s,i)ds \le \frac{N}{r} \frac{(2i)!}{(i!)^2 2^{2i+1}}$, and (3.26) now follows by Stirling's formula. From [12,15] we use the explicit representation of $e^{[A_1^N(t)]}$ given by

$$e^{[A_1^N t]} = \begin{pmatrix} I & 0 \\ \left[\exp(\frac{tN}{r}\, C^N) - I\right](C^N)^{-1} a^N & \exp(\frac{tN}{r}\, C^N) \end{pmatrix} \tag{3.27}$$

where $a^N = \text{col}(I,0,\ldots,0)$ and

$$C^N = \begin{pmatrix} -I & 0 & \cdots & 0 \\ I & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & 0 \\ 0 & \cdots\; 0 & I & -I \end{pmatrix}.$$

By employing the measure of the logarithmic norm of $[A_1^N]$ it is proved in [12, Lemma 5.4] that for each $x = (x_0,\ldots,x_N) \in \mathbb{R}^{n(N+1)}$ and with $\mathbb{R}^n$ endowed with the Euclidean norm

$$\| [e^{A_1^N t}] x \|^\infty \leq \|x\|^\infty . \tag{3.28}$$

holds; it is simple to check that endowing $\mathbb{R}^n$ with any norm does not change that estimate.

To verify (3.24) we recall the definition of $t^N_{j^N}$ and choose $k^N = k^N(\varepsilon)$ such that $\varepsilon - r_1 \in [t^N_{j^N-k^N}, t^N_{j^N-k^N-1})$; of course, it is still assumed that $\varepsilon < \varepsilon_1 < r_1$. The choice of k^N implies

$$\frac{(k^N-1)}{N} r \leq \varepsilon \leq \frac{(k^N+1)}{N} r \quad \text{and} \quad \left|\frac{\varepsilon}{k^N} - \frac{r}{N}\right| \leq \frac{r}{Nk^N} . \tag{3.29}$$

From (3.11), (3.28) and (3.29) it follows that

$$\|v^{N,\varepsilon}(t)\|^\infty \leq |P_1^N\phi| + \int_{t_0}^t |f(s,z^{N,\varepsilon}(s),u(s))|ds +$$

$$+ \left|\int_{t_0}^t \frac{1}{\varepsilon} B \operatorname{col}(I,(1-\sigma(1,t-s))I,\ldots,(1-\sigma(N,t-s))I)(v^{N,\varepsilon}_{j^N-k^N}(s)-v^{N,\varepsilon}_{j^N}(s))ds\right|$$

where $\sigma(j,s) = \exp(\frac{-Ns}{r}) \sum_{\lambda=0}^{j-1} (\frac{1}{\lambda !})(\frac{Ns}{r})^\lambda$ for $\lambda = 1,\ldots,N$. By (H3) the last estimate implies

$$\|v^{N,\varepsilon}\|^\infty \leq |P_1^N\phi| + \int_{t_0}^t n_2(s)(\|v^{N,\varepsilon}(s)\|^\infty + 1)(1+|u(s)|^{p/2})ds + \|R(t)\|^\infty,$$

where we yet have to find bounds on the coordinates of $R(t)$.

$$|R_0(t)| = \left|\frac{Br}{\varepsilon N}\int_{t_0}^{t}\sum_{\mu=j^N-k^N+1}^{j^N} \dot{v}_\mu^{N,\varepsilon}(s)ds\right| \leq$$

$$\leq \frac{\|B\|\, r}{\varepsilon N} k^N (\| v^{N,\varepsilon}(t)\|^\infty + \| v^{N,\varepsilon}(t_0)\|^\infty) \leq$$

$$\leq \|B\| (1+\frac{r}{\varepsilon N})(\| v^{N,\varepsilon}(t)\|^\infty + \| v^{N,\varepsilon}(t_0)\|^\infty);$$

here we used (3.29).

Similarly for $i = 1,\ldots,N$ one finds

$$|R_i(t)| \leq \left|\frac{Br}{\varepsilon N}\int_{t_0}^{t}\left[1-\exp(\frac{-N}{r}(t-s))\sum_{\lambda=0}^{i-1}\frac{1}{\lambda!}(\frac{N}{r})^\lambda(t-s)^\lambda\right]\sum_{\mu=k^N-j^N+1}^{j^N} \dot{v}_\mu^{N,\varepsilon}(s)ds\right| =$$

$$= \left|\frac{Br}{\varepsilon N}\sum_{\mu=k^N-j^N+1}^{j^N}\left[(\sigma(i,t-t_0)-1)v_\mu^{N,\varepsilon}(t_0)\right] + \int_{t_0}^{t} Z(s,i-1)v_\mu^{N,\varepsilon}(s)ds\right| \leq$$

$$\leq \frac{\|B\|\, rk^N}{\varepsilon N}\left[\|v^{N,\varepsilon}(t_0)\|^\infty + \int_{t_0}^{t} Z(s,i-1)\ v^{N,\varepsilon}(s)\|^\infty ds\right] \leq$$

$$\leq \|B\| (1+\frac{r}{\varepsilon N})\left(\|v^{N,\varepsilon}(t_0)\|^\infty + \int_{t_0}^{t} Z(s,i-1)\|v^{N,\varepsilon}(s)\|^\infty ds\right).$$

By hypothesis on ε and N, we have that $1 - \|B\| - \frac{r}{\varepsilon N} \geq \frac{1}{2}(1-\|B\|)$, and $\|B\|(1+\frac{r}{\varepsilon N}) \leq 1$, and therefore after summarizing the above estimates one arrives at

$$\|v^{N,\varepsilon}\|^\infty \leq 2(1-\|B\|)^{-1}\left[|P_1^N\phi| + \int_{t_0}^{t} n_2(s)(1+|u(s)|^{p/2})ds + \|v^{N,\varepsilon}(t_0)\|^\infty +\right.$$

$$+ \max_{i=1,\dots,N} \int_{t_0}^{t} (Z(s,i-1) + n_2(s)(1+|u(s)|^{p/2})) \|v^{N,\varepsilon}(s)\|^{\infty} ds \Big] \le$$

$$\le 2(1-\|B\|)^{-1}\Big[2|\phi| + \int_{t_0}^{t} n_2(s)(1+|u(s)|^{p/2})ds +$$

$$+ \max_{i=1,\dots N} \int_{t_0}^{t} (Z(s,i-1)+n_2(s)(1+|u(s)|^{p/2})) \|v^{N,\varepsilon}(s)\|^{\infty} ds\Big];$$

an application of Gronwall's inequality now implies the result. For the estimate of $\|\dot{v}^{N,\varepsilon}(t)\|^{\infty}$ one uses (3.11), (3.27) and (3.28), so that for $t \in [t_0,T]$

$$\|\dot{v}^{N,\varepsilon}(t)\|^{\infty} \le \|\tfrac{N}{r}\, \mathrm{col}(0,(v_0^{N,\varepsilon}(t_0) - v_1^{N,\varepsilon}(t_0))I,\dots,(v_{N-1}^{N,\varepsilon}(t_0) - v_N^{N,\varepsilon}(t_0))I\|^{\infty} +$$

$$+ \Big\|\int_{t_0}^{t} \mathrm{col}\Big(0,e^{\frac{-N(t-s)}{r}} I,\dots,\frac{1}{(N-1)!}\big(\tfrac{N(t-s)}{r}\big)^{N-1} e^{\frac{-N(t-s)}{r}} I\Big)$$

$$\Big[f(s,z^{N,\varepsilon}(s),u(s)) + \frac{1}{\varepsilon} B(v^{N,\varepsilon}_{j^N-k^N}(s) - v^{N,\varepsilon}_{j^N}(s))\Big]ds\Big\|^{\infty} +$$

$$+ |f(t,z^{N,\varepsilon}(t),u(t))| + |\frac{1}{\varepsilon} B(v^{N,\varepsilon}_{j^N-k^N}(t) - v^{N,\varepsilon}_{j^N}(t))|.$$

Recall that $P_1^N\phi = \sum_{j=0}^{N} \beta_j^N \phi(t_j^N) = \beta^N v^{N,\varepsilon}(t_0)$. Since by assumption $\phi \in W^{1,\infty}$, the last estimate together with (H3), (3.25) and the previously determined bound on $z^{N,\varepsilon}(t)$ (or equivalently $v^{N,\varepsilon}(t)$) imply

$$\|\dot{v}^{N,\varepsilon}(t)\|^{\infty} \le \text{ess sup}|\dot{\phi}| + \int_{t_0}^{t} n_2(s)(1+|u(s)|^{p/2})(1+K_5)ds + n_2(t)(1+|u(t)|^{p/2})(K_5+1) +$$

$$+ \left\|\int_{t_0}^{t} \mathrm{col}\left(0, e^{\frac{-N(t-s)}{r}} I, \ldots, \frac{1}{(N-1)!}\left(\frac{N(t-s)}{r}\right)^{N-1} e^{\frac{-N(t-s)}{r}} I\right)\frac{1}{\varepsilon}B(v^{N,\varepsilon}_{j^N-k^N}(s) - v^{N,\varepsilon}_{j^N}(s))ds\right\|^{\infty} +$$

$$+ \left|\frac{1}{\varepsilon} B(v^{N,\varepsilon}_{j^N-k^N}(t) - v^{N,\varepsilon}_{j^N}(t))\right|. \tag{3.30}$$

It is in this estimate that $\phi \in W^{1,\infty}$ is needed essentially.

By (3.10) the coefficient vectors $v^{N,\varepsilon}(t)$ satisfy

$$\dot{v}^{N,\varepsilon}(t) = [A_1^N]v^{N,\varepsilon}(t) + [Q_1^N]\{f(t,z^{N,\varepsilon}(t),u(t)) + \frac{1}{\varepsilon} B(v^{N,\varepsilon}_{j^N-k^N}(t) - v^{N,\varepsilon}_{j^N}(t))\}, \tag{3.31}$$

which implies that $\dot{v}_i^{N,\varepsilon}(t) = \frac{N}{r}(v_{i-1}^{N,\varepsilon}(t) - v_i^{N,\varepsilon}(t))$ for $i = 1,\ldots,N$. Therefore

$$\left|\frac{1}{\varepsilon} B(v^{N,\varepsilon}_{j^N-k^N}(t) - v^{N,\varepsilon}_{j^N}(t))\right| \le \frac{\|B\|\, r\, k^N}{\varepsilon N} \|\dot{v}^{N,\varepsilon}(t)\|^{\infty} \le \|B\|\, (1+\frac{r}{\varepsilon N}) \|\dot{v}^{N,\varepsilon}(t)\|^{\infty}.$$

The fourth term IV in (3.30) is estimated separately now. A short calculation shows

$$|(IV)_1| \le 3\frac{\|B\|\, r}{\varepsilon N} K_5 \le 3K_5 .$$

And for $i = 2,\ldots,N$ we get

$$|(IV)_i| = \left|\frac{Br}{\varepsilon N} \sum_{\mu = j^N-k^N+1}^{j^N} \int_{t_0}^{t} \frac{1}{(i-1)!}\left(\frac{N(t-s)}{r}\right)^{i-1} e^{\frac{-N(t-s)}{r}} \dot{v}^{N,\varepsilon}(s)ds\right| =$$

$$= \frac{\|B\| \, rk^N}{\varepsilon N} \sum_{\mu=j^N-k^N+1}^{j^N} \left[\|v^{N,\varepsilon}(t_0)\|^{\infty} + \int_{t_0}^{t} (Z(s,i-2) + Z(s,i-1)) \|v^{N,\varepsilon}(s)\|^{\infty} ds \right]$$

$$\leq \frac{3\|B\| \, rk^N K_5}{\varepsilon N} \leq 6K_5.$$

Using these last estimates in (3. 30) we arrive at

$$\|\dot{v}^{N,\varepsilon}(t)\|^{\infty} \leq 2(1-\|B\|)^{-1}\Big[6K_5 + |\dot{\phi}|_{\infty} + \int_{t_0}^{t} n_2(s)(1+|u(s)|^{p/2})(K_5+1)ds +$$

$$+ (K_5+1)\, n_2(t)(1+|u(t)|^{p/2})\Big],$$

so that (3.24) is established. The proof of (3.23), similar to the one for (3.24), is contained in [15] and will therefore not be given here.

Lemma 3.3. Under the hypotheses of Theorem 3.2 and the additional condition that $\frac{r}{\varepsilon N} \leq \frac{1-\|B\|}{2}$, there exist constants M_i depending on ϕ, f, B, T and U but independent of $\dot{\phi}, N, \varepsilon, t$ such that

$$|v_0^{N,\varepsilon}(t) - v_1^{N,\varepsilon}(t)| \leq |\dot{\phi}|_{\infty} M_1 \frac{1}{N} + M_2 \frac{1}{\sqrt{N}} \tag{3.32}$$

$$|w_0^{N}(t) - w_1^{N}(t)| \leq |\dot{\phi}|_{\infty} M_3 \frac{1}{N} + M_4 \frac{1}{\sqrt{N}} \tag{3.33}$$

$$\|v^{N,\varepsilon}(t) - w^{N}(t)\|^{\infty'} \leq (|\dot{\phi}|_{\infty} + 1)M_5\left(\varepsilon + \frac{1}{\sqrt{N}}\right) + M_6 \frac{1}{\varepsilon N}. \tag{3.34}$$

Proof. From (3.11), (3.24), (3.27) and (H3) it follows that

$$|v_0^{N,\varepsilon}(t) - v_1^{N,\varepsilon}(t)| \le |e^{\frac{-N}{r}(t-t_0)}((P_1^N\phi)_0 - (P_1^N\phi)_1)| +$$

$$+ |\int_{t_0}^{t} e^{\frac{-N(t-s)}{r}}(f(s,z^{N,\varepsilon}(s),u(s)) + \frac{1}{\varepsilon}B(v^{N,\varepsilon}_{j^N-k^N}(s) - v^{N,\varepsilon}_{j^N}(s)))| \le$$

$$\le |\phi(0) - \phi(t_1^N)| + \int_{t_0}^{t} e^{\frac{-N(t-s)}{r}} n_2(s)(1+|u(s)|^{p/2})(\|v^{N,\varepsilon}(s)\|^{\infty} + 1)ds +$$

$$+ |\frac{1}{\varepsilon} B\frac{r}{N} \sum_{\mu=j^N-k^N+1}^{j^N} \int_{t_0}^{t} e^{\frac{-N(t-s)}{r}} \dot{v}_{\mu}^{N,\varepsilon}(s)ds| \le$$

$$\le \frac{r}{N}|\dot{\phi}|_\infty + (K_5+1)n_2(t)\left(\frac{r}{N} + (\frac{r}{2N})^{1/2}\left(\int_{t_0}^{t}|u(s)|^p ds\right)^{1/2}\right) +$$

$$+ \frac{1}{\varepsilon}\|B\|\frac{r}{N}k^N\left[(|\dot{\phi}|_\infty K_6 + K_7)\frac{r}{N} + K_8(\frac{r}{2N})^{1/2}\left(\int_{t_0}^{t}|u(s)|^p ds\right)^{1/2}\right].$$

By (3.29) and since $\frac{r}{\varepsilon N} < 1$ it follows that $\frac{\|B\|\, rk^N}{\varepsilon N} < 2$, which can be used in the last estimate to imply (3.32). The proof of (3.33) is quite similar. Employing (3.26) the second estimate in (3.29) and (3.31) the verification of (3.34) is somewhat tedious, but simple, and will not be given here, see [16]. Estimates (3.32)-(3.34) imply (3.34) with $\|\cdot\|^{\infty'}$ replaced by $\|\cdot\|^{\infty}$; we were not able to prove (3.34) directly.

4. Application to Optimal Control Problems and Examples

In this section the theory that has been developed in the previous section will be applied to optimal control problems associated with (NFDE).

We restrict ourselves to a simple class of problems and refer to [1], [16] for a more elaborate discussion on the type of problems, specifically cost functionals, that can be treated within the same framework. The aim of this section is to demonstrate the feasibility of using the linear interpolating spline and the averaging approximation scheme for actual computations and to report on some examples.

The equation under consideration is (2.1) and (2.2) again and in view of Remark 3.3 we assume that $\mathbb{R}^n$ is endowed with a norm such that $\sum_{i=1}^{\ell} ||B_i|| < 1$. The controls will be taken from a closed and convex subset U of $L^p(t_0,T;\mathbb{R}^m)$. Below we shall continue to give the details for the $\{Z_1^N, P_1^N, A_1^N\}$ scheme and we only mention that in continuation of Remark 3.4 and Proposition 3.1 one may derive similar results for the averaging approximation scheme.

We define a functional $J: \mathbb{R}^n \times C \times C(t_0,T;\mathbb{R}^n) \times L^p(t_0,T;\mathbb{R}^m) \to \mathbb{R}$ by

$$J(y_1,y_2,y_3,y_4) = (y_1-\xi(0))^*G(y_1-\xi(0)) + \int_{-r}^{0} (y_2(s)-\xi(s))^*H(y_2(s)-\xi(s))ds + \\ + \int_{t_0}^{T} [y_3(s)^*Qy_3(s) + y_4(s)^*Ry_4(s)]ds \tag{4.1}$$

where $\xi \in C$ is fixed and G,H,Q and R are semi-definite, symmetric matrices of appropriate dimensions. R moreover is positive definite. The cost functional $\phi: L^p(t_0,T;\mathbb{R}^m) \to \mathbb{R}$ is given by

$$\phi(u) = J(x(T;\phi,u),\ x_T(\cdot;\phi,u),\ x(\cdot;\phi,u),u)$$

where $x(\cdot;\phi,u)$ is a solution of (2.1). Now an optimal control problem can be formulated:

(P) Minimize $\phi(u)$ over U.

Analogously we define a sequence of approximating optimal control problems:

(P^N) Minimize $\phi^N(u)$ over U, where

$$\phi^N(u) = J^N(z^N(T;\phi,u)(0),\ z^N(T;\phi,u),\ z^N(\cdot;\phi,u)(0),u),$$

with $z^N(\cdot;\phi,u)$ solution of (3.12) and with

$$J^N: \mathbb{R}^n \times C \times C(t_0,T;\mathbb{R}^n) \times L^p(t_0,T;\mathbb{R}^m) \to \mathbb{R},$$

given by

$$J^N(y_1,y_2,y_3,y_4) = (y_1-\xi(0))^*G(y_1-\xi(0)) + \int_{-r}^{0}(y_2(s) - (P_1^N\xi)(s))^*H(y_2(s) - (P_1^N\xi)(s))ds$$
$$+ \int_{t_0}^{T}[y_3(s)^*Qy_3(s) + y_4(s)Ry_4(s)]ds. \tag{4.2}$$

Note that (P^N) is a finite dimensional problem in the sense that it is equivalent to an ordinary differential equation optimal control problem for the coefficients $w^N(t)$ of $z^N(t)$. We shall not discuss the existence problem of solutions of (P^N) but rather refer to [e.g. [17], Chapters 4 and 5] and assume that (P^N) has a solution $\bar{u}^N$; if f is of the form

$$f(t,\phi,u) = \sum_{i=0}^{\nu} A_i\phi(-r_i) + Eu \tag{4.3}$$

with $A_i \in \mathscr{L}_{n,n}$, $E \in \mathscr{L}_{n,m}$ and $0 = r_0 \leq \ldots \leq r_\nu = r$, then $\bar{u}^N$ exists [16] and since $u \to \phi^N(u)$ is strictly convex it is the unique solution of (P^N).

<u>Theorem 4.1</u>. If (H1)-(H4) hold, if the norm on $\mathbb{R}^n$ is chosen such that $\sum_{i=1}^{\ell} ||B_i|| < 1$ for the subordinate matrix norms and if $\{\bar{u}^{N_k}\}$ is a sequence of optimal controls of (P^N) in the convex and closed subset U of $L^2(t_0,T;\mathbb{R}^m)$, then there exists a subsequence $\{\bar{u}^{N_k}\}$ of $\bar{u}^N$ which converges weakly to an optimal control $\bar{u} \in U$ of (P).

<u>Proof</u>. For the ease of the reader we shall include the proof; the arguments are quite standard, however. First, note that $\{\bar{u}^N\}$ is bounded in $L^2(t_0,T;\mathbb{R}^m)$. For if not, then there exists a subsequence $\bar{u}^{N_\ell}$ of $\bar{u}^N$ with $|\bar{u}^{N_\ell}| \to \infty$ and therefore $\phi^{N_\ell}(\bar{u}^{N_\ell}) \leq \phi^{N_\ell}(v) \to \phi(v) < \infty$ for any $v \in L^2(t_0,T;\mathbb{R}^m)$, which cannot hold for the specific J that was chosen. $\{\bar{u}^N\}$ being bounded implies that it is weakly compact. Therefore, there exists a subsequence

$\{\bar{u}^{N_k}\}$ of $\{\bar{u}^N\}$, which converges weakly to some $\bar{u}$ in $L^2(t_0,T;\mathbb{R}^m)$. U is assumed to be convex and closed, which implies that it is also weakly closed and therefore $\bar{u} \in U$. By the triangle inequality we find for $t \in [t_0,T]$

$$
\begin{aligned}
|x_t(\cdot;\phi,u) - z^{N_k}(t;\phi,\bar{u}^{N_k})| &\le |x_t(\cdot;\phi,u) - x_t(\cdot;\phi,\bar{u}^{N_k})| \\
&\quad + |x_t(\cdot;\phi,\bar{u}^{N_k}) - z^{N_k}(t;\phi,\bar{u}^{N_k})| .
\end{aligned}
\tag{4.4}
$$

By Lemma 2.3, the first term on the right hand side of (4.4) converges to 0 as $K \to \infty$. Convergence to zero for the second term is a consequence of Theorem 3.2. In the following estimate we use $u^{N_k} \rightharpoonup u$, the special form of J and (4.4) to find

$$
\begin{aligned}
\Phi(\bar{u}) &= J(x(T;\phi,\bar{u}),\ x_T(\cdot;\phi,\bar{u}),\ x(\cdot;\phi,\bar{u}),\bar{u}) = \\
&= \lim_{k\to\infty} J(z^{N_k}(T;\phi,u^{N_k})(0),\ z^{N_k}(T;\phi,\bar{u}^{N_k}),\ z^{N_k}(\cdot;\phi,\bar{u}^{N_k})(0),\ \bar{u}^{N_k}) = \\
&= \lim_{k\to\infty} J^{N_k}(z^{N_k}(T;\phi,u^{N_k})(0),\ z^{N_k}(T;\phi,\bar{u}^{N_k}),\ z^{N_k}(\cdot;\phi,\bar{u}^{N_k})(0),\ \bar{u}^{N_k}) = \\
&= \lim_{k\to\infty} \Phi^{N_k}(u^{N_k}) \le \overline{\lim}\, \Phi^{N_k}(u) = \Phi(u),
\end{aligned}
$$

for any $u \in U$. Therefore, $\bar{u}$ is indeed a solution of (P); moreover, if we put $u = \bar{u}$ in the above inequalities then it follows that $\lim_{k\to\infty} \Phi^{N_k}(u^{N_k}) = \Phi(u^*)$.

It was pointed out in [16, Remark 4.2] that for f of the form (4.4) the stronger result $\lim_N \bar{u}^N = \bar{u}$ in $L^2(t_0,T;\mathbb{R}^m)$ holds.

We have applied Theorem 4.1 to a number of examples of the type

$$
(P^*)\left\{\begin{array}{l}
\text{minimize} \\
\quad \tilde{\phi}(u) = \frac{1}{2}[x(T)^*Gx(T)] + \frac{1}{2}\int_{t_0}^{T} [x(t)^*Qx(t) + u(t)^*Ru(t)]dt \\
\quad \text{over } L^2(t_0,T;\mathbb{R}^m) \\
\text{where } G,Q,R \text{ are matrices of appropriate dimensions with} \\
G \geq 0,\ Q \geq 0,\ R > 0, \\
\text{subject to} \\
\quad \dot{x}(t) - B\dot{x}(t-r) = \hat{f}(t,x_t) + Eu(t),
\end{array}\right.
$$

here $\hat{f}: \mathbb{R} \times C \to \mathbb{R}^n$ is assumed to be continuously Fréchet differentiable in the second variable for each fixed t, $t \to \hat{f}(t,\varphi_t)$ is Borel measurable for each $\varphi \in C(t_0-r,T;\mathbb{R}^n)$ and further, given any compact convex set $X \subset \mathbb{R}^n$ there exists $m \in L^1(t_0,T;\mathbb{R}^1)$ such that $|\hat{f}(t,\varphi_t df(t,\chi)\psi| \leq m(t)$ and $|d\hat{f}(t,\varphi_t)(\cdot)| \leq m(t)$ for each $\varphi \in C(t_0-r,T;\mathbb{R}^n)$; here $d\hat{f}(t,\chi)\psi$ denotes the Fréchet derivative of $\hat{f}$ w.r.t. the second variable evaluated at ψ.

In some cases we can calculate analytical solutions to (P^*) using the maximum principle for NFDE. Since the latter is not readily available in the literature, we include its statement in a form modified to suit (P^*) in Proposition 4.1. This is a special case of a very general maximum principle for NFDE that was developed in [14, specifically Lemma 3.3 and Theorem 4.2]; see also [5].

Proposition 4.1. Let u^* be the optimal control for (P^*). Then there exists a scalar $\alpha_0 < 0$ and a function $\psi \in L^2(t_0,T+r;\mathbb{R}^n)$ such that

(i) $\psi(t) \equiv 0$ on $(T,T+r]$

(ii) $\psi(T)^* = \alpha_0 Gx(T)$ and

$$\psi(t) = \alpha_0 x(T)^* G + \psi(r+s)^* B - \int_t^T \psi(s)^* \eta(s,t-s)ds + \alpha_0 \int_t^T x(s)^* Qds$$

where $d\hat{f}(t,x_t(\bar{u})) = \int_{-r}^{0} d_s \eta(t,s)y(s)$ for $y \in C$, and

(iii) $$\int_{t_0}^{T} [(\psi(t))^* E\bar{u}(t) + \frac{1}{2}\alpha_0 \bar{u}(t)^* R\bar{u}(t)]dt =$$

$$= \max_{v \in L^p(t_0,T;\mathbb{R}^m)} \int_{t_0}^{T} [\psi(t)^* Ev(t) + \frac{1}{2}\alpha_0 v(t)^* Rv(t)]dt.$$

If $\hat{f}$ is moreover linear in the second variable then (i)-(iii) in Proposition 4.1 guarantee the existence of a solution $\bar{u}$ of (P^*), [14, Theorem 5.1, Remark 5.1] and (iii) can be replaced by the pointwise maximum principle. For

$$\hat{f}(t,\varphi) = A_0\varphi(0) + A_1\varphi(-r)$$

for example, we get the following necessary and sufficient conditions characterizing $\bar{u}$, (without loss of generality, we let $\alpha_0 = -1$):

(i)' $\psi(t) \equiv 0$ on $(T,T+r]$

(ii)' $$\psi(t) = -x(T)^* G + \psi(r+s)^* B + \int_t^T \psi(s)^* A_0 ds + \int_{t+r}^{T} \psi(s)^* A_1 ds - \int_t^T x(s)^* Qds$$

(iii)' $$\psi(t)^* E\bar{u}(t) - \frac{1}{2}\bar{u}(t)^* R\bar{u}(t)^* = \max_{v \in \mathbb{R}^m} \psi(t)^* Ev - \frac{1}{2} v^* Rv$$

for almost every $t \in [t_0,T]$.

From (iii)' it follows that $\bar{u}(t) = R^{-1}E^{*}\psi(t)$ and we see that even in this very simple case ψ, the solution of (ii)' and therefore $\bar{u}$ will <u>not</u> be continuous, in general, but will have jumps at all multiples values of r.

We shall now briefly report on some of the numerical examples that were carried out.

<u>Example 1</u>. This is the optimization problem of minimizing

$$J(u) = \frac{1}{2}\gamma x^2(s) + \frac{1}{2}\int_0^2 u^2(s)ds$$

over $L^2(0,2;\mathbb{R})$, subject to

$$\dot{x}(t) = \frac{1}{4}\dot{x}(t-1) + x(t-1) + u(t) \quad \text{for} \quad t \in [0,2]$$

$$x(t) = \alpha \quad \text{for} \quad t \in [-1,0].$$

We used the maximum principle in the form (i)'-(iii)' to calculate the exact solution. The optimal trajectory x_{ex} and optimal control u_{ex} were found to be

$$x_{ex}(t) = \begin{cases} \alpha t + \delta(\frac{t^2}{2} - \frac{9}{4}t) + 1 & \text{for} \quad t \in [0,1] \\ (1-\delta+\frac{\alpha}{4})t + \frac{(t-1)^2}{2}(\alpha - \frac{9\delta}{4}) + \frac{\delta}{8}(t-\frac{13}{4})^2 + \frac{\delta}{6}(t-1)^3 + \frac{3}{4}\alpha - \frac{177}{128}\delta & \text{for} \quad t \in [1,2] \end{cases}$$

and

$$u_{ex}(t) = \begin{cases} \delta(t - \frac{9}{4}) & \text{for} \quad t \in (0,1] \\ -\delta & \text{for} \quad t \in (1,2] \end{cases}$$

where $\delta = \dfrac{(2+\frac{7}{4}\alpha)\gamma}{1+\gamma\,\frac{199}{48}}$.

For this and all the other examples reported on in this section, we used the averaging scheme to approximate the infinite dimensional problem. The resulting finite dimensional optimum control problem was solved via a combined gradient-conjugate gradient iterative technique and numerical integration was carried out by a modified Runge-Kutta method (Gill's modification).

For this example, we did calculations for various values of α and γ. In Tables 1-3 the results for $\alpha \equiv 1$ and $\gamma = 1$ are presented. As should be expected one can recognize convergence of optimal state, payoff and control to the exact solutions. However, it seems quite difficult from this and the other examples that we studied on the computer to predict a possible rate-of-convergence result. Certainly, the convergence will be slower than linear, in general. On the other hand, in many examples we experienced surprisingly good approximation results for low values of N. As should be expected, due to the jumps in the optimal control at multiples of 1, the relative error of $u_{ex} - u^N$ has a peak at $t = 1$. One can convince oneself quickly that the fact that it actually increases in this example is no reason for precariousness.

TABLE I

t	$x_{ex}(t)$	$x_{ex}-x^4$	$x_{ex}-x^8$	$x_{ex}-x^{16}$	$x_{ex}-x^{32}$	$x_{ex}-x^{64}$	$x_{ex}-x^{128}$
0.0	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.25	0.8629	0.0118	0.0096	0.0078	0.0058	0.0040	0.0007
0.5	0.7713	0.0136	0.0077	0.0055	0.0039	0.0026	0.0017
0.75	0.7252	0.0214	0.0078	0.0023	0.0009	0.0008	0.0007
1.0	0.7247	0.0484	0.0261	0.0136	0.0070	0.0035	0.0018
1.25	0.7401	0.0734	0.0445	0.0262	0.0149	0.0082	0.0045
1.5	0.7384	0.0701	0.0384	0.0205	0.0112	0.0063	0.0000
1.75	0.7308	0.0567	0.0278	0.0136	0.0071	0.0040	0.0024
2.0	0.7287	0.0506	0.0257	0.0133	0.0070	0.0038	0.0021

TABLE II

t	$u_{ex}(t)$	$u_{ex}-u^4$	$u_{ex}-u^8$	$u_{ex}-u^{16}$	$u_{ex}-u^{32}$	$u_{ex}-u^{64}$	$u_{ex}-u^{128}$
0.0	-1.6397	0.0863	0.0748	0.0641	0.0539	0.0454	0.0390
0.25	-1.4575	0.0125	0.0078	0.0048	0.0008	-0.0023	-0.0032
0.5	-1.2753	-0.0237	-0.0250	-0.0195	-0.0119	-0.0067	-0.0038
0.75	-1.0931	-0.0268	-0.0318	-0.0291	-0.0199	-0.0098	-0.0039
1.0	-0.9109	0.0015	-0.0130	-0.0291	-0.0440	-0.0563	-0.0658
1.25	-0.7287	0.0644	0.0469	0.0256	0.0082	-0.0006	-0.0019
1.5	-0.7287	-0.0136	-0.0138	-0.0115	-0.0070	-0.0038	-0.0021
1.75	-0.7287	-0.0466	-0.0255	-0.0133	-0.0070	-0.0038	-0.0021
2.0	-0.7287	-0.0506	-0.0257	-0.0133	-0.0070	-0.0038	-0.0021

TABLE III

	exact	$J_{ex}-J^4$	$J_{ex}-J^8$	$J_{ex}-J^{16}$	$J_{ex}-J^{32}$	$J_{ex}-J^{64}$	$J_{ex}-J^{128}$
J	1.3664	0.0186	0.0081	0.0039	0.0021	0.0012	0.0008

Example 2. This is another linear optimization problem, in which we try to minimize

$$J(u) = \frac{1}{2}\gamma\, x^2(2) + \frac{1}{2}\int_0^2 \rho\, u^2(s)ds$$

over $u \in L^2(0,2;\mathbb{R})$, subject to

$$\dot{x}(t) = -\frac{1}{2}\dot{x}(t-1) + 2x(t) + 4u(t) \qquad \text{for} \quad 0 \le t \le 2$$

$$x(t) \equiv \alpha \qquad \text{for} \quad -1 \le t \le 0.$$

Again we used the maximum principle to calculate the exact solutions:

$$x_{ex}(t) = \begin{cases} \alpha e^{2t} + 16\rho^{-1}\delta e^{2(1+t)}[e^{-4t}(\frac{e^2}{4} + \frac{t}{4} - \frac{5}{16}) + \frac{5}{16} - \frac{e^2}{4}] & \text{for} \quad t \in [0,1] \\ \alpha e^{2t} + 4\rho^{-1}\delta e^{2t-2}(e^2 - e^6 + \frac{5e^4}{4} - \frac{1}{4}) + 4r^{-1}\delta(e^2 - \frac{5}{4})e^{2t}(t-1) - \alpha e^{2t-2}(t-1) + & \\ \quad + \delta\rho^{-1}[e^{4-2t}(\frac{13}{2} - t - e^2) + e^{2t}(e^2 - \frac{11}{2})] & \text{for} \quad t \in [1,2] \end{cases}$$

and

$$u_{ex}(t) = \begin{cases} 4\rho^{-1}\delta e^{2(1-t)}(\frac{3}{2} - e^2 - t) & \text{for} \quad t \in (0,1] \\ -4\rho^{-1}\delta e^{4-2t} & \text{for} \quad t \in (1,2], \end{cases}$$

where $\delta = \alpha(e^4 - e^2)(\gamma^{-1} - \rho^{-1}(-4e^8 + 10e^6 - \frac{13}{2}e^4 - 2e^2 + \frac{9}{2}))^{-1}.$

The exact and the numerical values for the values $A \equiv 4$, $\rho = 3$ and $\gamma = 1$ can be found in Tables 4-6 below.

The numerical solution of the optimal control problem requires an initial guess for the optimal control. In order to be able to compare between various examples we always take as an initial guess $u_0 \equiv 0$. In actual computations it will be advisable to use the solution u^N of the N^{th} approximative step as start-up value in the next step. The approximation in Example 2, for instance, is not as good as in Example 1 where the exact optimal control is closer to 0 from $t = 0$ on. This can be seen when comparing the relative errors $|\phi_{ex}-\phi^N|/|\phi_{ex}|$.

TABLE IV

t	$x_{ex}(t)$	$x_{ex}-x^4$	$x_{ex}-x^8$	$x_{ex}-x^{16}$	$x_{ex}-x^{32}$	$x_{ex}-x^{64}$	$x_{ex}-x^{128}$
0.0	4.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.25	2.3730	0.0247	0.0166	0.0101	0.0063	0.0040	0.0024
0.5	1.2449	0.0149	0.0313	0.0219	0.0105	0.0043	0.0014
0.75	0.3698	-0.1569	-0.0642	-0.0139	0.0049	0.0058	0.0019
1.0	-0.4503	-0.5966	-0.4510	-0.3358	-0.2492	-0.1836	-0.1342
1.25	-0.3738	-0.3618	-0.2246	-0.1176	-0.0544	-0.0267	-0.0174
1.5	-0.3019	-0.2629	-0.1737	-0.1095	-0.0725	-0.0503	-0.0355
1.75	-0.1851	-0.1877	-0.0514	-0.1287	-0.1097	-0.0974	-0.0642
2.0	0.0686	-0.0116	-0.0058	-0.0029	-0.0014	-0.0007	-0.0003

TABLE V

t	$x_{ex}(t)$	$u_{ex}-u^4$	$u_{ex}-u^8$	$u_{ex}-u^{16}$	$u_{ex}-u^{32}$	$u_{ex}-u^{64}$	$u_{ex}-u^{128}$
0.0	-3.9803	0.0119	0.0120	0.0122	0.0123	0.0122	0.0120
0.25	-2.5167	0.0320	0.0145	0.0049	-0.0001	-0.0024	-0.0030
0.5	-1.5886	0.0398	0.0183	0.0067	0.0013	-0.0004	-0.0009
0.75	-1.0012	0.0412	0.0222	0.0123	0.0063	0.0024	0.0004
1.0	-0.6302	0.0393	0.0250	0.0189	0.0168	0.0167	0.0175
1.25	-0.4099	0.0216	0.0106	0.0059	0.0043	0.0031	0.0018
1.5	-0.2486	0.0289	0.0175	0.0099	0.0051	0.0024	0.0011
1.75	-0.1508	0.0242	0.0127	0.0063	0.0031	0.0015	0.0007
2.0	-0.0915	0.0154	0.0007	0.0038	0.0018	0.0008	0.0004

TABLE VI

	exact	$J_{ex}-J^4$	$J_{ex}-J^8$	$J_{ex}-J^{16}$	$J_{ex}-J^{32}$	$J_{ex}-J^{64}$	$J_{ex}-J^{128}$
J	6.4798	-0.1895	-0.0935	-0.0431	-0.0177	-0.0054	0.0002

<u>Example 3</u>. This finally is the nonlinear example of minimizing

$$J(u) = x^2(2) + \frac{1}{2}\int_0^2 u^2(s)ds$$

over $u \in L^2(0,2;\mathbb{R})$ subject to

$$\dot{x}(t) = \sin x(t) + x(t-1) - \frac{1}{4}\dot{x}(t-1) + u(t) \quad \text{for} \quad t \in [0,2]$$

$$x(t) \equiv 4 \quad \text{for} \quad t \in [-1,0].$$

The numerical results of this example which are given in Tables 7-9 seem to support what could be seen in the linear examples already: the approximation is quite well for low values of N, and increasing N does not improve the accuracy very quickly.

TABLE VII

t	x^4-x^8	x^8-x^{16}	$x^{16}-x^{32}$	$x^{32}-x^{64}$	$x^{64}-x^{128}$	x^{128}
0.0	0.0000	0.0000	0.0000	0.0000	0.0000	4.0000
0.25	0.0213	0.0242	0.0167	0.0089	0.0038	3.8395
0.50	0.0266	0.0334	0.0271	0.0151	0.0066	3.7760
0.75	0.0037	0.0128	0.0160	0.0149	0.0109	3.8424
1.00	-0.0378	-0.0318	-0.0252	-0.0197	-0.0153	4.0646
1.25	-0.0609	-0.0476	-0.0312	-0.0166	-0.0078	4.1047
1.50	-0.0435	-0.0244	-0.0143	-0.0090	-0.0053	3.9466
1.75	-0.0158	-0.0048	0.0008	-0.0009	-0.0020	3.6203
2.00	-0.0057	0.0036	0.0068	0.0063	0.0046	3.1176

TABLE VIII

t	u^4-u^8	u^8-u^{16}	$u^{16}-u^{32}$	$u^{32}-u^{64}$	$u^{64}-u^{128}$	u^{128}
0.0	0.0977	0.0917	0.0650	0.0498	0.0426	-4.0700
0.25	0.0772	0.0982	0.0725	0.0285	0.0049	-3.8079
0.50	-0.0121	0.0103	0.0353	0.0412	0.0255	-3.4425
0.75	-0.1329	-0.1392	-0.1130	-0.0610	-0.0057	-2.7341
1.00	-0.1706	-0.1598	-0.1369	-0.1119	-0.0881	-2.4752
1.25	-0.0568	0.0135	0.0615	0.0565	0.0194	-3.4442
1.50	0.0598	0.0407	0.0044	-0.0049	-0.0037	-4.0264
1.75	0.0315	-0.0041	-0.0117	-0.0106	-0.0075	-4.9039
2.00	0.0118	-0.0067	-0.0137	-0.0126	-0.0093	-6.2344

TABLE IX

	J^4-J^8	J^8-J^{16}	$J^{16}-J^{32}$	$J^{32}-J^{64}$	$J^{64}-J^{128}$	J^{128}
J	0.0348	0.0541	0.0545	0.0428	0.0302	24.4971

Remark 4.1. A series of examples of linear and nonlinear unstable (see Remark 3.3) NFDE was also tested on the computer. The numerical results strongly indicate that the linear schemes presented in this paper do converge without assuming (HS). Unlike in

Examples 1-3, however, for low values of N the numerical solutions are worthless, their relative error being greater than 1 sometimes. Increasing N to 256 and higher, however, resulted in a surprising increase of the observed accuracy of the approximation. Although from a practical point of view the case of unstable neutral delay equations is not as important as the stable one, it should be a challenging question to study the piecewise linear schemes presented without assuming (H5). Let us recall that for linear NFDE and cubic (or higher-) order spline schemes (H5) could be avoided by working in state spaces endowed with a special norm depending on the equation, [9,11].

Acknowledgement. I would like to thank Professor H.T. Banks for making available to me the software packages that have been developed previously to study optimal control of functional differential equations and Patty Daniels for her assistance with the programs.

REFERENCES

[1] H.T. Banks, Approximation of nonlinear functional differential equation control systems, Journal of Optimization Theory and Applications, 29, (1979), 383-408.

[2] H.T. Banks and J.A. Burns, Hereditary control problems numerical methods based on averaging approximations, SIAM J. Control and Optimization 16,(1978), 169-208.

[3] H.T. Banks, J.A. Burns and E.M. Cliff, A comparison of numerical methods for identification and optimization problems involving control systems with delay, Brown University Technical Report #79-7.

[4] H.T. Banks and F. Kappel, Spline approximation for functional differential equations, J. Differential Eqs. 34(1979), 496-522.

[5] H.T. Banks and G.A. Kent, Control of functional differential equations of retarded and neutral type to target sets in function space, SIAM J. Control and Optimization 10(1972), 567-593.

[6] A. Halanay and V. Rasvan, Periodic and almost periodic solutions for a class of systems described by a coupled delay-differential and difference equations, J. Nonlinear Analysis: TMA, 1,(1977), 197-206.

[7] A. Halanay and V. Rasvan, Approximation of delays by ordinary differential equations, to appear in Proc. International Conf. Recent Adv. Diff. Eqs., Trieste, August 24-26, 1978.

[8] J.K. Hale, Theory of Functional Differential Equations, Springer-Verlag, 1977.

[9] F. Kappel, Approximation of neutral functional differential equations in the state space $\mathbb{R}^n \times L^2$, Colloquia Math. Soc. Janos Bolyai,30, Qualitative Theory of Diff. Eqns., to appear.

[10] F. Kappel, Differenzen-Differential gleichungen, 5. Steiermarkisches mathematisches Symposion, Stift Rein, Ber. math.-stat. Sekt. Forschungszentrum Graz, 1(1973), 1-77.

[11] F. Kappel and K. Kunisch, Spline approximations for neutral functional differential equations, submitted to SIAM J. of Numerical Analysis.

[12] F. Kappel and W. Schappacher, Autonomous nonlinear functional differential equations and averaging approximations, J. Nonlinear Analysis: TMA, 2(1978), 391-422.

[13] F. Kappel and W. Schappacher, Nonlinear functional differential equations and abstract integral equations, Proc. Royal Soc. Edinburgh, 84A,(1979), 71-92.

[14] G.A. Kent, Optimal control of functional differential equations of neutral type, Ph.D. Dissertation, Brown University, Providence, R.I. (1971).

[15] K. Kunisch, Neutral functional differential equations in L^p-spaces and averaging approximations, J. Nonlinear Analysis:TMA, 3(1979), 419-448.

[16] K. Kunisch, Approximation of optimal control problems for nonlinear hereditary systems of neutral type, to appear in Ber. math.-stat. Sekt. Forschungs Zentium Graz.

[17] E.B. Lee and L. Marcus, Foundations of Optimum Control Theory, Wiley & Sons (1967).

[18] W.C. Pickel, Numerical methods for time optimal control of linear systems with applications to neutral systems, Ph.D. Dissertation, University of Missouri, Columbia, Mo.

[19] P.M. Prenter, Splines and Variational Methods, Wiley & Sons (1975).

[20] G.I. Rosen, A discrete approximation framework for hereditary systems, Ph.D. Dissertation, Brown University, Providence, R.I. (1980).